



## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

(i) APPLICANT: Anderson, David J.  
Saito, Tetsuichiro

5 (ii) TITLE OF INVENTION: A NOVEL HOMEODOMAIN PROTEIN

(iii) NUMBER OF SEQUENCES: 22

(iv) CORRESPONDENCE ADDRESS:

10 (A) ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert  
(B) STREET: Four Embarcadero Center, Suite 3400  
(C) CITY: San Francisco  
(D) STATE: California  
(E) COUNTRY: United States  
(F) ZIP: 94111

15 (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

20 (vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/701,278  
(B) FILING DATE: 22-AUG-1996  
(C) CLASSIFICATION:

25 Q2 (viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Silva, Robin M.  
(B) REGISTRATION NUMBER: 38,304  
(C) REFERENCE/DOCKET NUMBER: A-63770-1

(ix) TELECOMMUNICATION INFORMATION:

30 (A) TELEPHONE: (415) 781-1989  
(B) TELEFAX: (415) 398-3249

## (2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

35 (A) LENGTH: 2424 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: unknown  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GCAGAGGTAG GCAGGGTTCC CGAGCCGCTC TCCCGGCTCC CTGCTCTGGG CCTTGGGGCT 60

40 CCACCGGCTT CTTGGCCCGA GCTGCTGCGC GTGCAGATGG CCTTGCGCGA TCGCCGGACC 120



CCGCTGCGGT GGCCAAGTGC AGGGCTTGTG GCTGGGACCC CTGAGAACCA GGAGCCAGAC 180  
 TGTGCTCAGC TTGCCAGGCC GGAGCCACGC ACGGGCACAA GTCTGTCAGG CCGCCATCAG 240  
 TCCTGGTCCA GCCGTCAGGG CCCATCCGAC CGTCGGCGAT GTTTTATTTT CACTGCCCCG 300  
 CACAGCTAGA GGGCACAGCG CCTTTTGGTA ACCACTCTAC GGGGGATTTT GATGATGGGT 360  
 5 TTCTTAGAAG AAAACAGCGC AGAAATCGGA CAACCTTCGC TCTTCAGCAG TTGGAAGCTC 420  
 TGGAGGCAGT CTTTGCCCAA ACACACTACC CAGATGTCTT CACCAGAGAA GAGCTAGCCA 480  
 TGAAAATAAA CCTCACAGAA GCCAGAGTGC AGGTTTGGTT CCAGAACCGA AGAGCCAAGT 540  
 GGAGGAAGAC AGAGAGAGGG GCCTCTGACC AGGAACCAGG GGCTAAGGAA CCCATGGCAG 600  
 AGGTGACACC ACCCCCAGTG AGGAACATCA ACTCTCCACC CCCAGGGGAC CAGGCCCGGG 660  
 10 GCAAGAAGGA GGCCCTGGAG GCCCAGCAGA GCCTGGGACG CACAGTGGGC CCCGCCGGGC 720  
 CTTTCTTCCC CTCCTGCTTG CCAGGGACCC TCCTGAACAC AGCCACTTAT GCCCAGGCCC 780  
 TGTCCCATGT GGCATCTCTG AAAGGGGGCC CACTGTGCTC TTGCTGCGTC CCAGACCCTA 840  
 TGGGGCTCTC CTTCCTCCCC ACTTACGGTT GCCAGAGTAA CCGCACAGCC AGCGTGGCTG 900  
 CCCTGCGCAT GAAGGCCCGC GAGCATTAG AAGCGGTCCT GCAGTCTGCC AACCTTCTGC 960  
 15 CGTCCACCAG CAGCAGCCCC GGCCCTGCCT CCAAGCAGGT GCCTCCAGAA GGCAGCCAGG 1020  
 ACAAGCCCTC CCCAACGAAG GAACAGAGCG AGGGAGAGAA GAGCGTATGA GGGTCCGGAG 1080  
 AACCAGCTG GGAGCCCTGC CCACCCCTGC TTCTCTCAGC CTCAGCCCTG CCAGCCTCTG 1140  
 AACCACAAGG AGTAGCCACC TCCTCATGGA TCTGACAGGG CAAACGGGAC CTGCAAGCTG 1200  
 GTTGAGACCT GAAGAGTCCC TCTAGAATTC TGCTGGTAGG CTGTGTTGTT CTCGCTTTTC 1260  
 20 CTTTGGTGAC ATTTTCCGAT GGCTCTTAGT GACTCTGGAC ACTGCTCTGT GATGAGGTCC 1320  
 CTGTTTTTTG CTTTTTGTTT TGTCTCTTTT TTTTGTGTTT GTTTTGTGTT ATTTTCCAGG 1380  
 CCAAGCAGCC TTGGAGCAAA GCAGATTAGT TTATTCCACC ATCCTTCTTG AGATATCTGG 1440  
 GAAGGTCTTG TCAATTCCAA GGAAGTGGC AAGGATCATC CGTGAAAGAT GCCAAGAAGT 1500  
 GACATCTCAT GACAGGAAAT GAGACGGGCA CTCCCATATT GCTTAAGAAC CACAGAACTG 1560  
 25 GTGGACTATC AGCCAGTTCT CACTCCCTTC AGCCAGGACT GGCATCGGCC TCCTTTGTCT 1620  
 TGTTTAAAGG AATTAGCTGA GGTTTTGGCT AGGAAGTGAC AAGATATGGG CTGAAGACAT 1680  
 TGTGGTCCTG ACCCTAGCAG ATCTCCCTGG GCACATCTGA CCTGGTCCAG TCAGGCAGGT 1740  
 TGTCAGTTCG GGGATGGGGG CTGCTCTGCT GATTCTGTGT GTGGGTTCCC TGCAATTAGA 1800



GTGTTCACTT GCAGGCCCG CTCTCTTCAG AAGAGTGATG GGAAGTTCAC CAATCAGAAT 1860  
 GTAGCTTTGT AGCCCAGGAA AGGACCAGAG TCCTTGAAGC GGTAGGAAAT CCCTAGGAAG 1920  
 GCCCCTTAAA TACTTATGCC CAGATGAGCT GCCCTTCTTC CTATCCCCGT ATGTCGAGAG 1980  
 GTTGACGAGA CAGGAAAGCC AGGAAGATGA CTCCGTGTGG CAGAAGAGAA TGGAGTCCAA 2040  
 5 AGGGCCAAC TTACAGAGA TTTCTGCCGC AGTTTAGCGT GGCTGTGTTC TTTCACGCGA 2100  
 TGGTGACTTC GGAGAGATCA GAGGGAGATG TGCAATAGCA TGAGCCCCGC TCCTGGCCCCG 2160  
 GGTCTGGAA AGGTTGTGGT TGTGGTGG CTTTGGCTGA TGATGTTTCC ACGCAAACAG 2220  
 ATATTGCTTT CATGATGGCT GTTCTCATTT CAGTTCTGAT AATCGAGACG CTGTGCTCCC 2280  
 AGGCGCTCTG CCTCCCCTTA ACTCTTCAGG AGCACCCCCT CCCCTGTAAT ACTCCTAAGT 2340  
 10 GTATCGTGCC TCACTTACGG TTA CTGCAAC ACATTTGATG GAACACACTG TCTCCTTTAA 2400  
 AAAAGAAAAA AAAAAAAAAA AAAA 2424

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- 15 (A) LENGTH: 263 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: unknown  
 (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

20 Met Phe Tyr Phe His Cys Pro Pro Gln Leu Glu Gly Thr Ala Pro Phe  
 1 5 10 15  
 Gly Asn His Ser Thr Gly Asp Phe Asp Asp Gly Phe Leu Arg Arg Lys  
 20 25 30  
 Gln Arg Arg Asn Arg Thr Thr Phe Ala Leu Gln Gln Leu Glu Ala Leu  
 25 35 40 45  
 Glu Ala Val Phe Ala Gln Thr His Tyr Pro Asp Val Phe Thr Arg Glu  
 50 55 60  
 Glu Leu Ala Met Lys Ile Asn Leu Thr Glu Ala Arg Val Gln Val Trp  
 65 70 75 80  
 30 Phe Gln Asn Arg Arg Ala Lys Trp Arg Lys Thr Glu Arg Gly Ala Ser  
 85 90 95  
 Asp Gln Glu Pro Gly Ala Lys Glu Pro Met Ala Glu Val Thr Pro Pro  
 100 105 110



Pro Val Arg Asn Ile Asn Ser Pro Pro Pro Gly Asp Gln Ala Arg Gly  
 115 120 125

Lys Lys Glu Ala Leu Glu Ala Gln Gln Ser Leu Gly Arg Thr Val Gly  
 130 135 140

5 Pro Ala Gly Pro Phe Phe Pro Ser Cys Leu Pro Gly Thr Leu Leu Asn  
 145 150 155 160

Thr Ala Thr Tyr Ala Gln Ala Leu Ser His Val Ala Ser Leu Lys Gly  
 165 170 175

10 Gly Pro Leu Cys Ser Cys Cys Val Pro Asp Pro Met Gly Leu Ser Phe  
 180 185 190

Leu Pro Thr Tyr Gly Cys Gln Ser Asn Arg Thr Ala Ser Val Ala Ala  
 195 200 205

Leu Arg Met Lys Ala Arg Glu His Ser Glu Ala Val Leu Gln Ser Ala  
 210 215 220

15 Asn Leu Leu Pro Ser Thr Ser Ser Ser Pro Gly Pro Ala Ser Lys Gln  
 225 230 235 240

Val Pro Pro Glu Gly Ser Gln Asp Lys Pro Ser Pro Thr Lys Glu Gln  
 245 250 255

20 Ser Glu Gly Glu Lys Ser Val  
 260

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 70 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: unknown  
 (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

30 Gly Phe Leu Arg Arg Lys Gln Arg Arg Asn Arg Thr Thr Phe Ala Leu  
 1 5 10 15

Gln Gln Leu Glu Ala Leu Glu Ala Val Phe Ala Gln Thr His Tyr Pro  
 20 25 30

Asp Val Phe Thr Arg Glu Glu Leu Ala Met Lys Ile Asn Leu Thr Glu  
 35 40 45

35 Ala Arg Val Gln Val Trp Phe Gln Asn Arg Arg Ala Lys Trp Arg Lys  
 50 55 60

Thr Glu Arg Gly Ala Ser  
 65 70



## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 70 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: unknown  
 (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Leu His Glu Lys Arg Lys Gln Arg Arg Ile Arg Thr Thr Phe Thr Ser  
 1 5 10 15  
 Ala Gln Leu Lys Glu Leu Glu Arg Val Phe Ala Glu Thr His Tyr Pro  
 20 25 30  
 Asp Ile Tyr Thr Arg Glu Glu Leu Ala Leu Lys Ile Asp Leu Thr Glu  
 35 40 45  
 Ala Arg Val Gln Val Trp Phe Gln Asn Arg Arg Ala Lys Phe Arg Lys  
 50 55 60  
 Gln Glu Arg Ala Ala Ser  
 65 70

## (2) INFORMATION FOR SEQ ID NO:5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 70 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: unknown  
 (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Leu Pro Leu Lys Arg Lys Gln Arg Arg Ser Arg Thr Thr Phe Thr Ala  
 1 5 10 15  
 Glu Gln Leu Glu Glu Leu Glu Arg Ala Phe Glu Arg Thr His Tyr Pro  
 20 25 30  
 Asp Ile Tyr Thr Arg Glu Glu Leu Ala Gln Arg Ala Lys Leu Thr Glu  
 35 40 45  
 Ala Arg Val Gln Val Trp Phe Ser Asn Arg Arg Ala Arg Trp Arg Lys  
 50 55 60  
 Gln Ala Gly Ala Asn Gln  
 65 70



## (2) INFORMATION FOR SEQ ID NO:6:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 70 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: unknown  
 (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Asn Val Ser Ser Ser Lys Lys Arg Arg His Arg Thr Thr Phe Thr Ser  
 1 5 10 15  
 Leu Gln Leu Glu Glu Leu Glu Lys Val Phe Gln Lys Thr His Tyr Pro  
 20 25 30  
 Asp Val Tyr Val Arg Glu Gln Leu Ala Leu Arg Thr Glu Leu Thr Glu  
 35 40 45  
 Ala Arg Val Gln Val Trp Phe Gln Asn Arg Arg Ala Lys Trp Arg Lys  
 50 55 60  
 Arg Glu Arg Tyr Gly Gln  
 65 70

## (2) INFORMATION FOR SEQ ID NO:7:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 70 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: unknown  
 (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Glu Lys Lys Lys Arg Lys Gln Arg Arg Asn Arg Thr Thr Phe Asn Ser  
 1 5 10 15  
 Ser Gln Leu Gln Ala Leu Glu Arg Val Phe Glu Arg Thr His Tyr Pro  
 20 25 30  
 Asp Ala Phe Val Arg Glu Asp Leu Ala Arg Arg Val Asn Leu Thr Glu  
 35 40 45  
 Ala Arg Val Gln Val Trp Phe Gln Asn Arg Arg Ala Lys Phe Arg Arg  
 50 55 60  
 Asn Glu Arg Ala Met Leu  
 65 70



## (2) INFORMATION FOR SEQ ID NO:8:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 70 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: unknown  
 (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

10 Thr Lys Arg Lys Lys Lys Gln Arg Arg Asn Arg Thr Thr Phe Asn Ser  
 1 5 10 15

Ser Gln Leu Gln Ala Leu Glu Arg Val Phe Glu Arg Thr His Tyr Pro  
 20 25 30

Asp Ala Phe Val Arg Glu Glu Leu Ala Arg Arg Val Asn Leu Ser Glu  
 35 40 45

15 Ala Arg Val Gln Val Trp Phe Gln Asn Arg Arg Ala Lys Phe Arg Arg  
 50 55 60

Asn Glu Arg Ala Met Leu  
 65 70

## (2) INFORMATION FOR SEQ ID NO:9:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 70 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: unknown  
 (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

20 Leu Gln Leu Lys Arg Lys Leu Gln Arg Asn Arg Thr Ser Phe Thr Gln  
 1 5 10 15

30 Glu Gln Ile Glu Ala Leu Glu Lys Glu Phe Glu Arg Thr His Tyr Pro  
 20 25 30

Asp Val Phe Ala Arg Glu Arg Leu Ala Ala Lys Ile Asp Leu Pro Glu  
 35 40 45

Ala Arg Ile Gln Val Trp Phe Ser Asn Arg Arg Ala Lys Trp Arg Arg  
 50 55 60

35 Glu Glu Lys Leu Arg Asn  
 65 70



## (2) INFORMATION FOR SEQ ID NO:10:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 70 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: unknown  
 (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

10 Tyr Ala Pro Lys Arg Lys Gln Arg Arg Tyr Arg Thr Thr Phe Thr Ser  
 1 5 10 15  
 Phe Gln Leu Glu Glu Leu Glu Lys Ala Phe Ser Arg Thr His Tyr Pro  
 20 25 30  
 Asp Val Phe Thr Arg Glu Glu Leu Ala Met Lys Ile Gly Leu Thr Glu  
 35 40 45  
 15 Ala Arg Ile Gln Val Trp Phe Gln Asn Arg Arg Ala Lys Trp Arg Lys  
 50 55 60  
 Gln Glu Lys Val Gly Pro  
 65 70

## (2) INFORMATION FOR SEQ ID NO:11:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 70 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: unknown  
 (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

20 Ile Ala Leu Lys Arg Lys Gln Arg Arg Cys Arg Thr Thr Phe Ser Ala  
 1 5 10 15  
 30 Ser Gln Leu Asp Glu Leu Glu Arg Ala Phe Glu Arg Thr Gln Tyr Pro  
 20 25 30  
 Asp Ile Tyr Thr Arg Glu Glu Leu Ala Gln Arg Thr Asn Leu Thr Glu  
 35 40 45  
 Ala Arg Ile Gln Val Trp Phe Ser Asn Arg Arg Ala Arg Leu Arg Lys  
 50 55 60  
 35 Gln His Thr Ser Val Ser  
 65 70



## (2) INFORMATION FOR SEQ ID NO:12:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 70 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: unknown  
 (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Asp Pro Asn Gly Ile Lys Lys Lys Lys Thr Arg Thr Thr Phe Thr Ala  
 1 5 10 15  
 Tyr Gln Leu Glu Glu Leu Glu Arg Ala Phe Glu Arg Ala Pro Tyr Pro  
 20 25 30  
 Asp Val Phe Ala Arg Glu Glu Leu Ala Ile Lys Leu Asn Leu Ser Glu  
 35 40 45  
 Ser Arg Val Gln Val Trp Phe Gln Asn Arg Arg Ala Lys Trp Arg Lys  
 50 55 60  
 His Glu Pro Pro Arg Lys  
 65 70

## (2) INFORMATION FOR SEQ ID NO:13:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 70 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: unknown  
 (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Gly Glu Ser Ala Ala Lys Arg Arg Arg Thr Arg Thr Asn Phe Ser Gly  
 1 5 10 15  
 Trp Gln Leu Glu Glu Leu Glu Ser Ala Phe Glu Ala Ser His Tyr Pro  
 20 25 30  
 Asp Val Phe Met Arg Glu Ala Leu Ala Met Arg Leu Asp Leu Leu Glu  
 35 40 45  
 Ser Arg Val Gln Val Trp Phe Gln Asn Arg Arg Ala Lys Trp Arg Lys  
 50 55 60  
 Arg Glu Gln Asn Arg Asn  
 65 70



## (2) INFORMATION FOR SEQ ID NO:14:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Phe Thr Ala Tyr Gln Leu Glu  
1 5

## (2) INFORMATION FOR SEQ ID NO:15:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Gln Val Trp Phe Gln Asn Arg  
1 5

## (2) INFORMATION FOR SEQ ID NO:16:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: DNA (genomic)

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 14..15
- (D) OTHER INFORMATION: /note= "The 'N' appearing at position 14 represents inosine."

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 17..18
- (D) OTHER INFORMATION: /note= "The 'N' appearing at position 17 represents inosine."



## (ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION: 26..27

(D) OTHER INFORMATION: /note= "The 'N' appearing at  
 5 location 26 represents inosine."

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CGGGATCCTT YACNGCNTAY CARYTNGA

28

## (2) INFORMATION FOR SEQ ID NO:17:

## (i) SEQUENCE CHARACTERISTICS:

10 (A) LENGTH: 27 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: unknown

(D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: DNA (genomic)

## 15 (ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION: 25..26

(D) OTHER INFORMATION: /note= "The 'N' appearing at  
 position 25 represents inosine."

## 20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CGGAATTCKR TTYTGRAACC ANACYTG

27

## (2) INFORMATION FOR SEQ ID NO:18:

## (i) SEQUENCE CHARACTERISTICS:

25 (A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: unknown

(D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

30 TCATGAAGTG TGACGTTGAC ATCC

24

## (2) INFORMATION FOR SEQ ID NO:19:

## (i) SEQUENCE CHARACTERISTICS:

35 (A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: unknown

(D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: DNA (genomic)



(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GTAAAACGCA GCTCAGTAAC AGTC

24

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 22 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: unknown  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

ACTGAAGCCA AGGTAGGATC CG

22

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- 15 (A) LENGTH: 24 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: unknown  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

20 CGGATCCTAC CTTGGCTTCA GTAG

24

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- 25 (A) LENGTH: 24 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: unknown  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CTACTGAAGC CAAGGTAGGA TCCG

24

*As Cancel*